Markov chain mixing time	Geometric bounds for eigenvalues of Markov chains
	Lecture 2-3
Lecturer: Michael Choi	*Scribe: Michael Choi*

1 Goal of this lecture

In previous lecture, we have seen that the rate of convergence of reversible Markov chain is determined by SLEM (second largest eigenvalue in modulus). In this lecture, we introduce modern techniques in bounding SLEM, and hence mixing time.

2 Preliminaries - graph theoretic concepts

Suppose that P is the transition matrix of a reversible ergodic Markov chain X. We first introduce a few graph-theoretic quantities induced by the transition graph of P. For all $i, j \in \mathcal{X}$, we define

$$Q(i,j) = Q(P)(i,j) := \pi(i)P(i,j).$$

Note that Q is symmetric since P is self-adjoint in $\ell^2(\pi)$. The transition graph induced by P has a vertex set equals to $\mathcal{X}$ and $e = (i,j)$ is an edge if $Q(e) = Q(i,j) \neq 0$. For any two distinct vertex $i, j \in \mathcal{X}$, we select only one path from i to j, that is a sequence $i, i_1, \ldots, i_m, j$ such that $P(i,i_1)P(i_1,i_2)\ldots P(i_m,j) > 0$. We pick the path such that a given edge appears at most once. Irreducibility of P guarantees such a path exists. Denote by $\Gamma = \Gamma(P)$ to be the collection of paths selected. For a path $\gamma_{ij} \in \Gamma$, we define

$$|\gamma_{ij}|_Q := \sum_{e \in \gamma_{ij}} \frac{1}{Q(e)}.$$

The Poincaré coefficient of P is then

$$\kappa(\Gamma) = \kappa(P,\Gamma) := \max_e \sum_{\gamma_{ij} \ni e} |\gamma_{ij}|_Q \pi(i)\pi(j).$$

For each $i \in \mathcal{X}$, we select one closed path $\sigma_i = \sigma_i(P)$ from i to i with an odd number of edges and any given edge appears at most once. Denote by $\Sigma = \Sigma(P)$ to be the collection of all such closed paths. For a path $\sigma_i \in \Sigma$, we define

$$|\sigma_i|_Q := \sum_{e \in \sigma_i} \frac{1}{Q(e)},$$

$$\alpha(\Sigma) = \alpha(P,\Sigma) := \max_e \sum_{\sigma_i \ni e} |\sigma_i|_Q \pi(i).$$

The final quantity of interest is the notion of conductance of P. For any subset $S \subseteq \mathcal{X}$, define

$$Q(S \times S^c) := \sum_{i \in S, j \in S^c} Q(i,j) = \sum_{i \in S, j \in S^c} \pi(i)P(i,j),$$

where S^c is the complement of S in $\mathcal{X}$. The conductance $h = h(P)$ is then given by

$$h = h(P) := \inf_{S;0<\pi(S)\leq 1/2} \frac{Q(S \times S^c)}{\pi(S)}.$$

3 Poincaré and Cheeger's inequality for bounding SLEM

Theorem 1 (Bounds on SLEM of P) *Suppose that P is a reversible ergodic transition matrix with stationary distribution π and eigenvalues $1 = \lambda_1(P) > \lambda_2(P) \geq \ldots \geq \lambda_{|\mathcal{X}|}(P)$. Then we have*

1. *(Poincaré inequality for P)*

$$\lambda_2(P) \leq 1 - \frac{1}{\kappa(P,\Gamma)}.$$

2. *(Cheeger's inequality for P)*

$$1 - 2h(P) \leq \lambda_2(P) \leq 1 - \frac{h(P)^2}{2}.$$

3. *(Lower bound on $\lambda_{|\mathcal{X}|}(P)$)*

$$\lambda_{|\mathcal{X}|}(P) \geq -1 + \frac{2}{\alpha(P,\Sigma)}.$$

3.1 Proof of Theorem 1 item (1) Poincaré inequality

We follow the proof of (Diaconis and Stroock, 1991, Proposition 1) and (Brémaud, 1999, Theorem 4.1). We note that, for $f \in \ell^2(\pi)$,

$$\begin{aligned}
\mathrm{Var}_\pi(f) &= \frac{1}{2}\sum_{i,j}(f(i)-f(j))^2\pi(i)\pi(j) \\
&= \frac{1}{2}\sum_{i,j}\left(\sum_{e=(e^-,e^+)\in\gamma_{ij}}\sqrt{\frac{Q(e)}{Q(e)}}(f(e^-)-f(e^+))\right)^2\pi(i)\pi(j) \\
&\leq \frac{1}{2}\sum_{i,j}|\gamma_{ij}|_Q\left(\sum_{e=(e^-,e^+)\in\gamma_{ij}}Q(e)(f(e^-)-f(e^+))^2\right)\pi(i)\pi(j) \\
&= \frac{1}{2}\sum_{e=(e^-,e^+)}Q(e)(f(e^-)-f(e^+))^2\sum_{\gamma_{ij}\ni e}|\gamma_{ij}|_Q\pi(i)\pi(j) \leq \kappa(P,\Gamma)\langle(I-P)f,f\rangle_\pi,
\end{aligned}$$

where the inequality follows from Cauchy-Schwarz inequality. Upon rearranging and taking infimum over non-trivial f with mean 0 under stationarity (i.e. $\mathbb{E}_\pi f = 0$), desired result follows. That is,

$$\frac{1}{\kappa(P,\Gamma)} \leq \inf_{f\neq 0:\mathbb{E}_\pi f=0}\frac{\langle(I-P)f,f\rangle_\pi}{\mathrm{Var}_\pi(f)} = 1-\lambda_2(P).$$

3.2 Proof of Theorem 1 item (3) lower bound on $\lambda_{|\mathcal{X}|}(P)$

We follow the proof of (Diaconis and Stroock, 1991, Proposition 2) and (Brémaud, 1999, Theorem 4.2). First, we note that

$$\frac{1}{2}\sum_{i,j}(f(i)+f(j))^2 Q(i,j) = \langle Pf, f\rangle_\pi + \langle f, f\rangle_\pi.$$

If σ_i is a path of the form $(i_0 = i, i_1, i_2, \ldots, i_{2m}, i)$ with odd number of edges, then

$$\begin{aligned} f(i) &= \frac{1}{2}\left((f(i_0)+f(i_1)) - (f(i_1)+f(i_2)) + \ldots + (f(i_{2m})+f(i))\right) \\ &= \frac{1}{2}\sum_{e=(e^-,e^+)\in\sigma_i} (-1)^{n(e)}(f(e^-)+f(e^+)), \end{aligned}$$

where $n(e) = k$ if $e = (i_k, i_{k+1})$. As a result, using Cauchy-Schwarz inequality again we have

$$\begin{aligned} \langle f, f\rangle_\pi &= \sum_i \frac{\pi(i)}{4}\left(\sum_{e=(e^-,e^+)\in\sigma_i}\sqrt{\frac{Q(e)}{Q(e)}}(-1)^{n(e)}(f(e^-)+f(e^+))\right)^2 \\ &\leq \sum_i \frac{\pi(i)}{4}|\sigma_i|_Q\left(\sum_{e=(e^-,e^+)\in\sigma_i} Q(e)(f(e^-)+f(e^+))^2\right) \\ &= \frac{1}{4}\sum_{e=(e^-,e^+)} Q(e)(f(e^-)+f(e^+))^2\left(\sum_{\sigma_i\ni e}|\sigma_i|_{|Q|}\pi(i)\right) \\ &\leq \frac{\alpha(P,\Sigma)}{4}\sum_{e=(e^-,e^+)} Q(e)(f(e^-)+f(e^+))^2 \\ &= \frac{\alpha(P,\Sigma)}{2}\left(\langle Pf, f\rangle_\pi + \langle f, f\rangle_\pi\right). \end{aligned}$$

Desired result follows from the variational characterization of the smallest eigenvalue of P.

3.3 Proof of Theorem 1 item (2) Cheeger's inequality for P

We follow the proof of (Brémaud, 1999, Theorem 4.3). We first prove the lower bound of item (2). Using the variational characterization of $\lambda_2(P)$, for any $f \in \ell^2(\pi)$ with $f \neq 0$ and $\mathbb{E}_\pi f = \langle \mathbf{1}, f\rangle_\pi = 0$, we have

$$1 - \lambda_2(P) \leq \frac{\langle (I-P)f, f\rangle_\pi}{\langle f, f\rangle_\pi}.$$

For any subset $S \subset \mathcal{X}$ such that $0 < \pi(S) \leq 1/2$, we take f to be

$$f(i) = \begin{cases} 1-\pi(S), & \text{if } i \in S; \\ -\pi(S), & \text{if } i \notin S. \end{cases}$$

It can be checked that $\langle \mathbf{1}, f\rangle_\pi = 0$ and $\langle f, f\rangle_\pi = \pi(S)(1-\pi(S))$. In addition,

$$\begin{aligned}\langle (I-P)f, f\rangle_\pi = \frac{1}{2}\sum_{i,j} Q(i,j)(f(j)-f(i))^2 &= \frac{1}{2}\sum_{i\in S}\sum_{j\in S^c} Q(i,j) + \frac{1}{2}\sum_{i\in S^c}\sum_{j\in S} Q(i,j)\\ &= \frac{1}{2}Q(S\times S^c) + \frac{1}{2}Q(S^c\times S)\\ &= Q(S\times S^c),\end{aligned}$$

where the last equality follows from $Q(i,j) = Q(j,i)$ as P is self-adjoint in $\ell^2(\pi)$. Collecting the above results give

$$1-\lambda_2(P) \le \frac{Q(S\times S^c)}{\pi(S)(1-\pi(S))} \le 2\frac{Q(S\times S^c)}{\pi(S)}.$$

Since this holds for any subset $S \subset \mathcal{X}$ such that $0 < \pi(S) \le 1/2$, minimizing over all such S gives the desired result.

Next, we prove the upper bound in item (2). Let u be a left eigenvector of P associated with eigenvalue $\lambda \neq 1$, and note that u is orthogonal to π, the eigenvector associated with 1. As a result u must have both positive and negative entries, and so does f defined via

$$f(i) := \frac{u(i)}{\pi(i)}.$$

Assume without loss of generality that for some $k \in \{1,2,\ldots,|\mathcal{X}|\}$, we have

$$f(1) \ge f(2) \ge \ldots \ge f(k) > 0 \ge f(k+1)\ldots \ge f(|\mathcal{X}|),$$

and that for $S := \{1,2,\ldots,k\}$ we have $\pi(S) \le 1/2$. This can be done by changing the order of the states. For the second assumption, we can change f into $-f$ if necessary. Now, define y to be

$$y(i) := \frac{u(i)}{\pi(i)}\mathbb{1}_{\{u(i)>0\}},$$

where $\mathbb{1}_A$ is the indicator function of the set A. Note that

$$u^T(I-P)y = (1-\lambda)u^Ty = (1-\lambda)\sum_{i\in S} y(i)^2\pi(i) = (1-\lambda)\langle y, y\rangle_\pi. \tag{1}$$

On the other hand, by definition we see

$$\begin{aligned}u^T(I-P)y &= \sum_{i\in S}\sum_{j\in\mathcal{X}}(\mathbb{1}_{\{i=j\}} - P(j,i))u(j)y(i)\\ &\ge \sum_{i\in S}\sum_{j\in S}(\mathbb{1}_{\{i=j\}} - P(j,i))u(j)y(i)\\ &= \langle y, (I-P)y\rangle_\pi = \sum_{i<j} Q(i,j)(y(i)-y(j))^2,\end{aligned} \tag{2}$$

where in the inequality we observe the missing terms $-P(j,i)f(j)y(i)$ for $i \in S$ and $j \in S^c$ are non-negative. For $a, b \in \mathbb{R}$, using $(a+b)^2 \le 2(a^2+b^2)$, together with the symmetry of Q and $P(i,j) \le 1$ for all $i \ne j$, we have

$$\begin{aligned}\sum_{i<j} Q(i,j)(y(i)+y(j))^2 &\le 2\sum_{i<j} Q(i,j)(y(i)^2+y(j)^2) \\ &= 2\left(\sum_{i<j} Q(i,j)y(i)^2 + \sum_{i<j} Q(j,i)y(j)^2\right) \\ &= 2\sum_{i\ne j} \pi(i)P(i,j)y(i)^2 \le 2\langle y,y\rangle_\pi \end{aligned} \tag{3}$$

Collecting the above equations (1), (2) and (3) yield

$$\begin{aligned}1-\lambda &\ge \frac{\sum_{i<j} Q(i,j)(y(i)-y(j))^2}{\langle y,y\rangle_\pi}\frac{\sum_{i<j} Q(i,j)(y(i)+y(j))^2}{2\langle y,y\rangle_\pi} \\ &\ge \frac{1}{2}\left(\frac{\sum_{i<j} Q(i,j)(y(i)^2-y(j)^2)}{\langle y,y\rangle_\pi}\right)^2. \end{aligned} \tag{4}$$

Define $S_l := \{1, 2, \ldots, l\}$. We calculate

$$\begin{aligned}\sum_{i<j} Q(i,j)(y(i)^2-y(j)^2) &= \sum_{i<j} Q(i,j) \sum_{i\le l<j} (y(l)^2-y(l+1)^2) \\ &= \sum_{l=1}^{k}(y(l)^2-y(l+1)^2)\sum_{i\in S_l, j\notin S_l} \pi(i)P(i,j) \\ &= \sum_{l=1}^{k}(y(l)^2-y(l+1)^2)Q(S_l\times S_l^c) \\ &\ge h(P)\sum_{l=1}^{k}(y(l)^2-y(l+1)^2)\pi(S_l) \\ &= h(P)\sum_{l=1}^{k}(y(l)^2-y(l+1)^2)\sum_{i=1}^{l}\pi(i) \\ &= h(P)\sum_{i=1}^{k}\pi(i)\sum_{l=i}^{k}(y(l)^2-y(l+1)^2) \\ &= h(P)\sum_{i=1}^{k}\pi(i)y(i)^2, \end{aligned} \tag{5}$$

where the inequality comes from the definition of $h(P)$ with $0 < \pi(S_l) \le \pi(S) \le 1/2$. Desired result can be obtained via collecting (4) and (5).

4 Example: random walk on graph

Consider an undirected graph $G = (V, E)$, where V is the set of vertices and E is the set of edges. We write $y \sim x$ if y is a neighbour of x. A simple random walk on the graph G with state space $\mathcal{X} = V$ has transition matrix given by

$$P(x, y) = \begin{cases} \frac{1}{d_x}, & \text{if } y \sim x, \\ 0, & \text{otherwise .} \end{cases}$$

The stationary distribution is $\pi(x) = \frac{d_x}{2|E|}$, and it can be checked that the random walk is reversible with respect to π. Let $d := \max_x d_x$ be the maximum degree, $|\gamma|_Q := \max_{i,j} |\gamma_{ij}|_Q$ and

$$B := \max_e |\{\gamma \in \Gamma;\ e \in \gamma\}|.$$

Then

$$\kappa(\Gamma) = \max_e \frac{1}{2|E|} \sum_{\gamma_{ij} \ni e} |\gamma_{ij}| d_i d_j \leq \frac{1}{2|E|} |\gamma|_Q d^2 B,$$

$$\lambda_2 \leq 1 - \frac{2|E|}{d^2 |\gamma|_Q B}.$$

Similar calculations give

$$\lambda_{|\mathcal{X}|} \geq -1 + \frac{2}{d|\sigma|b},$$

where $|\sigma| := \max_i |\sigma_i|$ and $b := \max_e |\{\sigma \in \Sigma;\ e \in \sigma\}|$.

References

P. Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Gibbs fields, Monte Carlo simulation, and queues.

P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.